On the asymptotic behaviour of random recursive trees in random environment

K.A. Borovkov[*], V.A. Vatutin[†]

February 2, 2008

Abstract

We consider growing random recursive trees in random environment, in which at each step a new vertex is attached (by an edge of a random length) to an existing tree vertex according to a probability distribution that assigns the tree vertices masses proportional to their random weights. The main aim of the paper is to study the asymptotic behaviour of the distance from the newly inserted vertex to the tree's root and that of the mean numbers of outgoing vertices as the number of steps tends to infinity. Most of the results are obtained under the assumption that the random weights have a product form with independent identically distributed factors.

Keywords: random recursive trees, random environment, Sptizer's condition, distance to the root, outdegrees.

AMS classification: primary 05C80, secondary 60G50, 05C05, 60F99.

1 Introduction

We consider the following random recursive tree model. A recursive tree is constructed incrementally, by attaching a new vertex to a randomly chosen existing tree vertex at each step. Initially, the tree consists of a single vertex $v(0)$ that has weight $w(0) = 1$ and label 0. At the first step, a new vertex $v(1)$ is added to the tree as a child of the initial vertex. It is labelled by 1, and a random weight $w(1) > 0$ and a random length $Y(1) \geq 0$ are assigned to the vertex and to the edge connecting the vertices $v(0)$ and $v(1)$, respectively. It is assumed that the edge is directed from $v(0)$ to $v(1)$. At step $j > 1$, given all the weights $w(0), w(1), \ldots, w(j-1)$, first a node $v(j^*)$ is chosen at random from the nodes $v(0), v(1), \ldots, v(j-1)$ according to the distribution with probabilities

[*]Department of Mathematics and Statistics, The University of Melbourne, Parkville 3010, Australia

[†]Steklov Mathematical Institute RAS, Gubkin St. 8, 119991 Moscow, Russia

proportional to the nodes' weights, and then a new vertex $v(j)$ is added to the tree as a child of the node $v(j^*)$. The new vertex has label j, and a random weight $w(j) > 0$ and a random length $Y(j) \geq 0$ are assigned to the vertex $v(j)$ and to the edge, connecting the vertices $v(j^*)$ and $v(j)$, respectively. As at the initial step (where, for consistency, we will put $1^* = 0$), the edge is directed from $v(j^*)$ to its child vertex $v(j)$. We assume that $\{Y(j)\}_{j\geq 1}$ is a sequence of independent random variables (r.v.'s) which is independent of the sequence of the (generally speaking, random) weights $\{w(j)\}_{j\geq 0}$. Interpreting the sequence of weights as a "random environment" in which our recursive tree is growing, and appealing to an analogy with random walks and branching processes in random environments, it is not unnatural to refer to such a model as a random recursive tree in random environment.

Let

$$D_0 := 0, \qquad D_n := D_{n^*} + Y(n), \qquad n \geq 1,$$

be the distance from the vertex $v(n)$ to the root (i.e. the sum of the lengths of the edges connecting $v(n)$ with $v(0)$). In this paper, we study the asymptotic (as $n \to \infty$) behavior of D_n under various assumptions on the random weights $w(j)$ and lengths $Y(j)$, and also that of the mean values of the outgoing degrees

$$N_n(j) := \sum_{k=j+1}^{n} I\{v(k^*) = v(j)\}, \qquad j \leq n, \tag{1}$$

where $I\{A\}$ is the indicator of the event A.

Observe that if $w(j) \equiv Y(j) \equiv 1$ for all j, then we get the standard random recursive tree ([11]; see also [16]). If $w(j) = a^j$, $j \geq 0$, where $a > 0$ is a constant and $Y(j)$, $j \geq 1$, are r.v.'s whose distributions satisfy certain mild conditions, we get the recursive tree considered in [10] (in fact, the model in [10] assumed that, at each step, a fixed number $k \geq 1$ of children are attached to one of the existing tree vertices, and also that $Y(j)$ are vector-valued).

One should also mention here other related models where the weights of the vertices can *change* at each step. Thus, if, after the completion of the kth step of the tree construction, the weight of the vertex $v(j)$, $j \leq k$, is $w(j) = w(j,k) = 1 + \beta N_k(j)$, where $\beta \geq 0$ is constant and $Y(j) \equiv 1$, we get the linear recursive tree studied in [17, 5] (see also the bibliography there for further references). The case when $w(j) = w(j,k) = 1 + N_k(j)$ was considered in [4]; the power-tail limiting behavior of the degree distribution for this model that had been guessed in [4] was established in [8].

If $w(j) = a_1 \cdots a_j$, $j \geq 1$, where $a_1, \ldots, a_j$ are independent and identically distributed (i.i.d) r.v.'s, and $Y(j) \equiv 1$, we get a version of a weighted recursive tree. It is this last model and its generalizations that will be of the main interest for us in the present paper.

From now on we assume that the weight $w(j)$ of the vertex $v(j)$ is, generally speaking, random and, once assigned, remains unchanged forever.

Section 2 of the paper is devoted to studying the asymptotic behaviour of the distribution of D_n. Theorems 1 and 2 present general convergence results for

the conditional distribution of D_n in the cases when the random weights $w(j)$ tend to "prescribe" new attachments to vertices close to the root of the tree and when the new attachments are "more dispersed" across the tree, respectively. Corollary 2 covers the special case when $w(j) \equiv 1$. The results of the section also show that, for any $\alpha \in (0, 1]$, one can construct a random recursive tree such that D_n behaves as n^α as $n \to \infty$. Theorem 3 implies that, in the case of the "product-form" weights $w(j) = a_1 \cdots a_j$, $j \geq 1$, with a_j being non-degenerate i.i.d. satisfying the moment conditions $\mathbf{E} \ln a_j = 0$ and $\mathbf{E} |\ln a_j|^{2+\delta} < \infty$ for a $\delta > 0$, the limiting distribution of $D_n/\sqrt{n}$ coincides with the law of the maximum of the Brownian motion process on a finite time interval.

Section 3 deals with the expectations of the numbers of outgoing degrees in the case of the product-form weights under the assumption that the random walk generated by the i.i.d. sequence $\{\ln a_j\}$ satisfies Spitzer's condition. Theorem 4 gives the asymptotic behaviour of the unconditional expectations $\mathbf{E} N_n(k)$ as $n \to \infty$ when either $k = j$ or $k = n - j$ for a fixed value $j \geq 0$ (in both cases it is given by a regularly varying function of n). Theorem 5 complements it by covering the case when $\min\{j, n - j\} \to \infty$ (here the answer has the form of a product of regularly varying functions of j and $n - j$, respectively; in particular, in the case when $\ln a_j$ has zero mean and a finite variance, one obtains $\mathbf{E} N_n(j) \sim 2\pi^{-1}(n-j)^{1/2} j^{-1/2}$). Theorem 6 describes, in a range of j-values, the asymptotic behaviour of the distribution of the conditional expectation $\mathbf{E}_w N_n(j)$ given the sequence of the weights $w(1), w(2), \ldots$.

2 The distribution of D_n

2.1 The basic properties of D_n

Let

$$W_n := \sum_{j=0}^{n} w(j), \qquad p_n(j) := \frac{w(j)}{W_n}, \quad j = 0, 1, \ldots, n.$$

Set $f_0(t) := 1$, $f_j(t) := \mathbf{E}\, e^{itY(j)}$, $j \geq 1$, and put $\varphi_0(t) := 1$,

$$\begin{aligned} \varphi_n(t) &:= \mathbf{E}_w e^{itD_n} := \mathbf{E}\left[e^{itD_n} \,\middle|\, w(1), \ldots, w(n-1)\right], \\ \Psi_n(t) &:= \mathbf{E}\, \varphi_n(t) = \mathbf{E}\, e^{itD_n}, \qquad n \geq 1 \end{aligned}$$

(here and in what follows, $\mathbf{E}_w$ and $\mathbf{P}_w$ denote the conditional expectation and probability given the sequence of weights $\{w(j)\}$, respectively).

It is easy to see that

$$
\begin{aligned}
\varphi_{n+1}(t) &= \sum_{j=0}^{n} p_n(j)\varphi_j(t) f_{n+1}(t) \\
&= \frac{W_{n-1}}{W_n} \sum_{j=0}^{n-1} p_{n-1}(j)\varphi_j(t) f_{n+1}(t) + p_n(n)\varphi_n(t) f_{n+1}(t) \\
&= (1 - p_n(n)) \frac{f_{n+1}(t)}{f_n(t)} \varphi_n(t) + p_n(n)\varphi_n(t) f_{n+1}(t) \\
&= \big[1 + (f_n(t) - 1)p_n(n)\big] \frac{f_{n+1}(t)}{f_n(t)} \varphi_n(t) = \cdots \\
&= f_{n+1}(t) \prod_{j=1}^{n} \big[1 + (f_j(t) - 1)p_j(j)\big]. \qquad (2)
\end{aligned}
$$

Remark 1. Observe that (2) means in fact that, given the environment, the r.v. D_{n+1} admits a representation in the form of a sum of independent r.v.'s as follows:

$$D_{n+1} \stackrel{d}{=} I_1 Y(1) + \cdots + I_n Y(n) + Y(n+1), \qquad (3)$$

where $\{I_j\}$ is a sequence of independent (of each other and also of $\{Y(j)\}$) random indicators with $\mathbf{P}(I_j = 1) = p_j(j)$, $j \geq 1$. In the special case when $Y(j) \equiv w(j) \equiv 1$, this representation is equivalent to the correspondence between the quantity D_n and the numbers of records in an i.i.d. sequence that was used in [11] (see also Section 3.6 in [18] for a discussion of a somewhat more general situation where the representation (3) with $Y(j) \equiv 1$ holds). Note, however, that in [11] a probabilistic argument that works in that special case only was used to derive the representation (3) which is actually the main tool for studying D_n, whereas our approach leads directly to (3) and is much more general.

From the recursive relation (2) one can derive a number of interesting results on the limiting behaviour of D_n. Note that (2) was first derived in the case when $w(j) = a^j$, $j \geq 0$, $Y(j) \in \mathbb{R}^d$, in [10] (one can easily see that this recursive formula and the statements of Theorems 1–2 below remain true in the multivariate case as well).

In particular, the relation (2) immediately implies the following assertion, describing the limiting behaviour of the conditional (given the weights) distribution of D_n when the weight sequence $\{w(j)\}$ "suggests" new children to attach not too far from the tree's root.

Theorem 1. *If*

$$\sum_{j=1}^{\infty} p_j(j) < \infty \qquad a.s.$$

and the distribution of $Y(n)$ *has a weak limit as* $n \to \infty$:

$$\lim_{n\to\infty} f_n(t) = f(t),$$

then there exists the limit

$$\lim_{n\to\infty}\varphi_n(t)=\varphi_\infty(t):=f(t)\prod_{j=1}^{\infty}\big[1+(f_j(t)-1)p_j(j)\big] \qquad a.s.$$

This result, in turn, implies that $D_n \xrightarrow{d} D_\infty$ as $n\to\infty$, where D_∞ is a proper r.v. with the characteristic function $\mathbf{E}\,\varphi_\infty(t)$.

The next assertion refers to situations where the attachment preferences are spread "more uniformly" across the tree.

Theorem 2. *Let the sequence of r.v.'s* $\{Y(j)\}_{j\ge1}$ *be uniformly integrable, and let there exist a sequence* $h_n\to\infty$, $n\to\infty$, *and a r.v.* ζ *such that the following convergence in distribution takes place as* $n\to\infty$:

$$\zeta_n := \frac{1}{h_n}\sum_{j=1}^{n} p_j(j)\mathbf{E}\,Y(j) \xrightarrow{d} \zeta. \tag{4}$$

Then for any t

$$\varphi_n\left(\frac{t}{h_n}\right) \xrightarrow{d} e^{it\zeta}.$$

Remark 2. One can easily see that if, instead of (4), one has $\zeta_n\to\zeta$ a.s. for some r.v. ζ, then

$$\lim_{n\to\infty}\varphi_n\left(\frac{t}{h_n}\right)=e^{it\zeta} \qquad \text{a.s.}$$

uniformly in t from any compact set.

Proof. It is not difficult to see that, due to the uniform integrability condition, as $n\to\infty$,

$$f_j\left(\frac{t}{h_n}\right)-1=\mathbf{E}\,\exp\left\{\frac{itY(j)}{h_n}\right\}-1=\frac{1}{h_n}\big(it\mathbf{E}\,Y(j)+o(1)\big)$$

uniformly in $j\ge1$ and in t from any compact set. Hence, as $p_j(j)\le1$, we have by (2)

$$\varphi_n\left(\frac{t}{h_n}\right)=f_n\left(\frac{t}{h_n}\right)\prod_{j=1}^{n-1}\left[1+\left(f_j\left(\frac{t}{h_n}\right)-1\right)p_j(j)\right]$$
$$=(1+\varepsilon_n(t))\exp\left\{\frac{it}{h_n}\sum_{j=1}^{n}p_j(j)\mathbf{E}\,Y(j)\right\},$$

where $\varepsilon_n(t)=o_P(1)$ as $n\to\infty$. This clearly implies the assertion of the theorem. $\square$

Corollary 1. *Under the conditions of Theorem 2,*

$$\lim_{n\to\infty} \Psi_n\left(\frac{t}{h_n}\right) = \mathbf{E}\, e^{it\zeta},$$

so that $\dfrac{D_n}{h_n} \xrightarrow{d} \zeta$ *as* $n \to \infty$*.*

From Theorem 2 one can also easily deduce the following result obtained in [10] (note that in the special case when $Y(j) \equiv 1$ the result was originally established in [11]).

Corollary 2. *If* $w(j) \equiv 1$, $j = 0, 1, 2, \ldots$, *the family of r.v.'s* $\{Y(j)\}_{j\ge 1}$ *is uniformly integrable and, as* $n \to \infty$,

$$\frac{1}{n}\sum_{j=1}^{n} \mathbf{E}\, Y(j) \to \mu \in \mathbb{R},$$

then $\dfrac{D_n}{\ln n} \xrightarrow{p} \mu$*.*

Proof. In this case clearly $p_j(j) = 1/(j+1)$, and, as it was shown in Lemma 1(i) in [10], under the above conditions

$$\zeta_n = \frac{1}{\ln n}\sum_{j=1}^{n} \frac{1}{j+1}\, \mathbf{E}\, Y(j) \to \mu.$$

Now the assertion of the theorem follows from Theorem 2. □

We also get the same asymptotics for D_n when the weights are random, but remain "on the average" the same.

Corollary 3. *If* $Y(j) \equiv 1$, $j \ge 1$, *and the sequence of random weights* $\{w(j)\}$ *satisfies the strong law of large numbers: as* $n \to \infty$,

$$\frac{1}{n}\sum_{j=1}^{n} w(j) \to a > 0 \quad \textit{a.s.},$$

then $\dfrac{D_n}{\ln n} \xrightarrow{p} 1$*.*

Proof. It again suffices to apply (a slightly modified version) of Lemma 1(i) from [10] (this time to the sequences $y_n := anW_n^{-1}$, $x_n := w(n)/a$) and use our Theorem 2. □

Remark 3. In fact, to obtain a faster than logarithmic growth rate for D_n (assuming that $Y(j) \equiv 1$), the weights $w(j)$ should grow faster that any power function. Indeed, if, say,

$$w(j) = j^{\alpha} l(j), \qquad \alpha \in \mathbb{R},$$

is a regularly varying function, then for $\alpha < -1$ one clearly has

$$\sum_{j=1}^{\infty} p_j(j) < \infty$$

(so that in this case Theorem 1 is applicable), whereas for $\alpha > -1$ by Karamata's theorem $W_n \sim (\alpha+1)^{-1} n^{\alpha+1} l(n)$, so that in this case $p_j(j) \sim 1/(\alpha+1)j$ and hence

$$\sum_{j=1}^{n} p_j(j) \sim \frac{\ln n}{\alpha+1}.$$

Thus, in the latter case $\dfrac{D_n}{\ln n} \xrightarrow{p} \dfrac{1}{\alpha+1}$ as $n \to \infty$.

On the other hand, for, say,

$$w(j) = \alpha j^{\alpha-1} e^{j^\alpha}, \qquad \alpha \in (0,1],$$

we get $W_n \sim e^{n^\alpha}$ and hence

$$\sum_{j=1}^{n} p_j(j) \sim n^\alpha.$$

So this example shows that, for any $\alpha \in (0,1]$, one can construct a random recursive tree with $\dfrac{D_n}{n^\alpha} \xrightarrow{p} 1$ as $n \to \infty$.

2.2 The case of the product-form random weights

In this subsection we will construct and study recursive trees with random vertex weights of the form $w(j) = a_1 \cdots a_j$, $j \geq 1$, where a_j are i.i.d. r.v.'s, and unit edge lengths. As it will be clearly seen from the proofs below, the main results will still hold true in the case of random i.i.d. edge lengths $Y(j) \geq 0$ with a finite mean as well (Remark 4). Thus restricting our attention to the case of unit edge lengths leads to no loss of generality, but makes the exposition more compact and transparent.

Denote by $\mathcal{T}_n$, $n = 0, 1, 2, \ldots$, the set of all rooted recursive trees having n nonrooted vertices and unit edge lengths (that is, $\mathcal{T}_n$ consists of the rooted trees whose root is labelled by 0 and whose nonrooted vertices are labelled by numbers $1, 2, \ldots, n$ in such a way that for any nonrooted vertex labelled, say, by j, the shortest path leading from it to the root traverses only the vertices whose labels are less than j). For a tree $t_n \in \mathcal{T}_n$, let $t_n(j) \in \mathcal{T}_{n+1}$ be the recursive tree which is obtained from t_n by adding a vertex labelled by $n+1$ as a child of the vertex with the label $j \in \{0, 1, \ldots, n\}$.

One can describe the construction of our random recursive tree as follows. First, we run a random walk

$$S_0 = 0, \qquad S_j = \theta_1 + \cdots + \theta_j, \quad j \geq 1,$$

where $\theta_j \stackrel{d}{=} \theta$, $j = 1, 2, \ldots, n$, are i.i.d. r.v.'s. Second, given S_j, $j = 0, 1, \ldots, n$, we construct a (conditional) Markov chain $T_0, T_1, \ldots, T_n$ with $T_k \in \mathcal{T}_k$, $k = 0, 1, \ldots, n$, by assigning the weight

$$w(j) := e^{-S_j}$$

to the vertex labelled by $j \geq 0$ (so that $w(j) = a_1 \cdots a_j$, $j \geq 1$, in the notation of Section 1, with $a_j := e^{-\theta_j}$ being i.i.d. r.v.'s), so that now we have, for $r = 0, 1, \ldots, n$,

$$W_r \equiv \sum_{q=0}^{r} w(q) = \sum_{q=0}^{r} e^{-S_q}, \tag{5}$$

$$p_r(j) \equiv \frac{e^{-S_j}}{W_r} = \frac{e^{-S_j}}{\sum_{q=0}^{r} e^{-S_q}}, \quad j = 0, 1, \ldots, r, \tag{6}$$

and then letting, for any $t_r \in \mathcal{T}_r$,

$$\begin{aligned} &\mathbf{P}_w\big(T_{r+1} = t_r(j) \big| \, T_r = t_r\big) \\ &\qquad \equiv \mathbf{P}\left(T_{r+1} = t_r(j) \big| \, T_r = t_r;\, w(0), w(1), \ldots, w(r)\right) := p_r(j), \end{aligned}$$

$j = 0, 1, \ldots, r$.

The main result of this subsection is

Theorem 3. *If*

$$\mathbf{E}\,\theta = 0, \qquad \sigma^2 := \mathbf{E}\,\theta^2 > 0, \qquad \mathbf{E}\,|\theta|^{2+\delta} < \infty \quad \textit{for a} \;\; \delta > 0, \tag{7}$$

then, as $n \to \infty$,

$$\zeta_n := \frac{1}{\sqrt{n}} \sum_{j=1}^{n} p_j(j) \xrightarrow{d} \sigma_m \max_{0 \leq u \leq 1} B(u), \qquad \sigma_m := \sigma \int_0^\infty \frac{m(dy)}{y} < \infty,$$

where $\{B(u)\}_{u \geq 0}$ *is the standard Brownian motion process and the measure* m *is specified in the proof* (see (11)).

Together with Corollary 1, the above assertion immediately yields the following

Corollary 4. *Under the conditions of Theorem 3*,

$$\frac{D_n}{\sqrt{n}} \xrightarrow{d} \sigma_m \max_{0 \leq u \leq 1} B(u) \quad \textit{as} \quad n \to \infty.$$

In other words, for any $x > 0$

$$\mathbf{P}\left(D_n > \sigma_m \sqrt{n} x\right) \to 2(1 - \Phi(x)),$$

where Φ *is the standard normal distribution function.*

Remark 4. It is obvious that the assertion of the corollary remains true in the case of i.i.d. random edge lengths $Y(j) \geq 0$ with a finite mean, with the only difference that σ_m should be replaced in its formulation with $\sigma_m \mathbf{E}\, Y(1)$.

Proof. Put

$$L_n := \min_{0 \leq k \leq n} S_k.$$

Basing on the proof of Theorem 4.1 in [3], we will show that

$$\frac{1}{|L_n|} \sum_{j=1}^{n} p_j(j) \to \int_0^\infty \frac{m(dy)}{y} < \infty \qquad \text{a.s.} \tag{8}$$

Since by the invariance principle

$$\frac{|L_n|}{\sqrt{n}} \xrightarrow{d} \sigma \max_{0 \leq u \leq 1} B(u) \qquad \text{as} \quad n \to \infty, \tag{9}$$

the assertion of the theorem will then immediately follow from (8).

First denote by

$$\gamma_0 := 0, \qquad \gamma_{j+1} := \min\{n > \gamma_j : S_n < S_{\gamma_j}\}, \quad j \geq 0,$$

the strict descending ladder epochs of the random walk $\{S_n\}_{n \geq 0}$. All the r.v.'s introduced are finite a.s. as $\{S_n\}_{n \geq 0}$ is recurrent in view of (7).

Let $\{X_n\}_{n \geq 0}$ be a Markov chain defined for $n = 1, 2, \dots$ by

$$X_n := e^{\theta_n} X_{n-1} + 1.$$

When $X_0^x = x > 0$ is a fixed value, we will use notation $\{X_n^x\}_{n \geq 0}$. Clearly,

$$X_n^x = x e^{S_n} + \sum_{q=1}^{n} e^{S_n - S_q} = e^{S_n}(x - 1 + W_n). \tag{10}$$

Set $\gamma := \gamma_1$. Under our assumptions (7), the expectation $\mathbf{E}\, S_\gamma < 0$ is finite (see e.g. Corollary 10, § 17 in [9]), and the Markov chain $\{X_{\gamma_n}\}_{n \geq 1}$ with the transition kernel

$$M_\gamma(x, A) := \mathbf{P}\,(X_\gamma^x \in A), \qquad x > 0, \quad A \in \mathcal{B},$$

has a unique invariant probability measure m_γ (see e.g. Lemma 5.49 in [13] and p.481 in [3]):

$$m_\gamma(A) = \int_0^\infty m_\gamma(dx) M_\gamma(x, A).$$

Moreover, the measure m defined by

$$m(f) := \frac{1}{\mathbf{E}\,(-S_\gamma)} \int_0^\infty \mathbf{E} \left(\sum_{k=0}^{\gamma-1} f(X_k^x) \right) m_\gamma(dx) \tag{11}$$

is an invariant measure for the Markov chain $\{X_n\}_{n\geq 0}$ (see [3]).

Now note that, by virtue of (6) and (10),

$$\zeta_n = \frac{1}{\sqrt{n}}\sum_{j=1}^{n} p_j(j) = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\frac{1}{X_j^1}\,.$$

Let P_{δ_y} be the distribution of the two-dimensional random walk

$$Z_n := (X_n, e^{S_n}), \quad n \geq 0$$

(on the group of transformations $x \mapsto ax+b$ of the real line with the composition law $(b_1, a_1)(b_2, a_2) = (b_1 + a_1 b_2, a_1 a_2)$) when $X_0 = y$. It was shown in the proof of Theorem 4.1 of [3] that if $f \in L^1(m)$ then

$$\lim_{n\to\infty}\frac{1}{|L_n|}\sum_{j=1}^{n} f(X_j) = \int_0^\infty f(y)\, m(dy) \qquad P_{m_\gamma}\text{-a.s.}, \tag{12}$$

where

$$P_{m_\gamma} := \int_0^\infty P_{\delta_y}\, m_\gamma(dy)$$

is the law of the two-dimensional random walk $\{Z_n\}_{n\geq 1}$ when the distribution of X_0 is m_γ.

Let for $N = 1, 2, \ldots$ and $x > 0$

$$g_N(x) := \frac{1}{x} I\{N^{-1} \leq x \leq N\} \leq \frac{1}{x} =: g(x).$$

Clearly, for all $x > 0$

$$g_N(x) \nearrow g(x) \quad \text{ as } \; N \to \infty, \tag{13}$$

and $g_N(x) \in L^1(m)$ for each $N = 1, 2, \ldots$. Therefore by (12)

$$\lim_{n\to\infty}\frac{1}{|L_n|}\sum_{j=1}^{n} g_N(X_j) = \int_0^\infty g_N(y)\, m(dy) \qquad P_{m_\gamma}\text{-a.s.} \tag{14}$$

On the other hand, for each $N \geq 1$ and any $x > 0$

$$\begin{aligned}
\frac{1}{|L_n|}\sum_{j=1}^{n} g_N(X_j^x) &\leq \frac{1}{|L_n|}\sum_{j=1}^{n} g(X_j^x) = \frac{1}{|L_n|}\sum_{j=1}^{n}\frac{e^{-S_j}}{x-1+W_j} \\
&\leq \frac{1}{|L_n|}\sum_{j=1}^{n}\int_{x-1+W_{j-1}}^{x-1+W_j}\frac{dy}{y} = \frac{1}{|L_n|}\int_{x-1+W_0}^{x-1+W_n}\frac{dy}{y} \\
&= \frac{1}{|L_n|}\left[\ln(x-1+W_n) - \ln x\right] \leq \frac{1}{|L_n|}\left[\ln\bigl(x+ne^{|L_n|}\bigr) - \ln x\right] \\
&\leq \frac{1}{|L_n|}\left[\ln ne^{|L_n|} + \frac{x}{ne^{|L_n|}} - \ln x\right] = 1 + \frac{1}{|L_n|}\bigl[\ln n + O(1)\bigr] \xrightarrow{p} 1
\end{aligned} \tag{15}$$

as $n \to \infty$ by the invariance principle (see e.g. [6]).

Combining (14) with (15) shows that

$$\sup_{N \geq 1} \int_0^\infty g_N(y)\, m(dy) \leq 1,$$

which together with (13) yields

$$\int_0^\infty g(y)\, m(dy) \leq 1.$$

Therefore by (12)

$$\lim_{n\to\infty} \frac{1}{|L_n|} \sum_{j=1}^n g(X_j) = \int_0^\infty g(y)\, m(dy) = \int_0^\infty \frac{dm(y)}{y} \qquad P_{m_\gamma}\text{-a.s.} \tag{16}$$

To see that this convergence holds for all starting points $x > 0$, it suffices to observe that $g(z)$ is monotone in $z > 0$ and $X_j^{x_1} > X_j^{x_2}$, $j \geq 1$, for $x_1 > x_2 > 0$.

This, in view of (8) and (9), completes the proof of Theorem 3.

□

3 The expectations of the outdegrees of vertices

Let $N_n(j)$ be the outdegree of the vertex $v(j)$, $j = 0, 1, \ldots, n$, in $\mathcal{T}_n$, i.e. the number of the edges coming out of $v(j)$ in a tree having n nonrooted vertices. Clearly, the r.v. $N_n(j)$ admits the representation (1), and therefore

$$\begin{aligned}\mathbf{E}_w N_n(j) &= \mathbf{E}\left[N_n(j) \big| \, w(1), \ldots, w(n-1)\right] \\ &= \sum_{k=j+1}^n \mathbf{E}_w I\{v(k^*) = v(j)\} = \sum_{k=j+1}^n p_{k-1}(j) = e^{-S_j} \sum_{k=j}^{n-1} W_k^{-1} \end{aligned} \tag{17}$$

and

$$\mathbf{E}\, N_n(j) = \sum_{k=j}^{n-1} \mathbf{E}\, e^{-S_j} W_k^{-1}. \tag{18}$$

Our aim in this section is to investigate the asymptotic (as $n \to \infty$) behavior of the expectations $\mathbf{E}\, N_n(j)$ and that of the distributions of the r.v.'s $\mathbf{E}_w N_n(j)$ in different ranges of the parameter j values.

3.1 The asymptotic behavior of $\mathbf{E}\, N_n(j)$

In this section we impose weaker restrictions (compared to the conditions (7) used in Section 2) on the random walk $S_n = \theta_1 + \cdots + \theta_n$, $n \geq 1$, where $\theta_j \stackrel{d}{=} \theta$ are i.i.d. r.v.'s. Namely, we only assume that Spitzer's condition holds:

There exists a $\rho \in (0,1)$ *such that*

$$\frac{1}{n}\sum_{k=1}^{n}\mathbf{P}\left(S_k>0\right)\to\rho \qquad as \quad n\to\infty. \tag{19}$$

It is known [12] that this condition is equivalent to Doney's condition

$$\mathbf{P}\left(S_n>0\right)\to\rho \qquad \text{as} \quad n\to\infty \tag{20}$$

(for a further discussion of the condition (19), see e.g. Section 8.9 in [7]).

We will need a number of auxiliary results concerning the random walk $\{S_n\}_{n\geq 0}$. Let

$$\Gamma_0:=0, \qquad \Gamma_{j+1}:=\inf\{n>\Gamma_j:\ S_n>S_{\Gamma_j}\}, \quad j\geq 0,$$

be the strict ascending ladder epochs of the random walk $\{S_n\}_{n\geq 0}$. Recall that $0=\gamma_0<\gamma_1<\gamma_2<\ldots$ denote the strict descending ladder epochs in the walk. Introduce the two renewal functions

$$U(x):=1+\sum_{j=1}^{\infty}\mathbf{P}\left(S_{\Gamma_j}<x\right), \quad x>0; \quad U(0)=1, \qquad U(x)=0, \quad x<0,$$

$$V(x):=\sum_{j=0}^{\infty}\mathbf{P}\left(S_{\gamma_j}\geq -x\right), \qquad x>0; \quad V(0)=1, \qquad V(x)=0, \quad x<0,$$

and set

$$M_n:=\max_{0\leq k\leq n}S_k, \qquad \widetilde{M}_n:=\max_{1\leq k\leq n}S_k.$$

It is known (see e.g. Lemma 1 in [15] and Lemma 1 in [20]) that under the condition (19)

$$\mathbf{E}\,U(-\theta)I\{-\theta>0\}=e^{-\phi}, \quad \mathbf{E}\,U(x-\theta)I\{x-\theta>0\}=U(x), \quad x>0, \tag{21}$$

where

$$\phi:=\sum_{j=1}^{\infty}\frac{1}{j}\mathbf{P}\left(S_j=0\right)<\infty, \tag{22}$$

and

$$\mathbf{E}\,V(x+\theta)=V(x), \quad x\geq 0. \tag{23}$$

By means of $V(x)$ and $U(x)$ one can specify two sequences of probability measures $\{\mathbf{P}_n^-\}_{n\geq 1}$ and $\{\mathbf{P}_n^+\}_{n\geq 1}$ on the σ-algebras $\{\Sigma_n:=\sigma(S_1,\ldots,S_n)\}_{n\geq 1}$, respectively, with the corresponding expectations $\{\mathbf{E}_n^-\}_{n\geq 1}$ and $\{\mathbf{E}_n^+\}_{n\geq 1}$, by setting for each bounded measurable function $\psi_n(x_1,\ldots,x_n)$

$$\mathbf{E}_n^-\left[\psi_n(S_1,\ldots,S_n)\right]:=e^{\phi}\mathbf{E}\left[\psi_n(S_1,\ldots,S_n)U(-S_n)I\{\widetilde{M}_n<0\}\right] \tag{24}$$

and

$$\mathbf{E}_n^+\left[\psi_n(S_1,\ldots,S_n)\right]:=\mathbf{E}\left[\psi_n(S_1,\ldots,S_n)V(S_n)I\{L_n\geq 0\}\right]. \tag{25}$$

It is easy to verify that (21) and (23) imply that each of the sequences $\{\mathbf{P}_n^{\pm}\}_{n\ge1}$ is consistent, and therefore by Kolmogorov's extension theorem there exist measures $\mathbf{P}^-$ and $\mathbf{P}^+$ on the σ-algebra $\sigma(S_1,S_2,\dots)$ such that their restrictions $\mathbf{P}^{\pm}|_{\Sigma_n}$ to Σ_n coincide with $\mathbf{P}_n^{\pm}, n=1,2,\dots$.

It is known (see Lemma 2.7 in [1]) that, under the condition (19),

$$\eta_1 := \sum_{k=1}^{\infty} e^{S_k} < \infty \qquad \mathbf{P}^-\text{-a.s.}, \qquad \eta_2 := \sum_{k=0}^{\infty} e^{-S_k} < \infty \qquad \mathbf{P}^+\text{-a.s.} \tag{26}$$

Finally, it is not difficult to deduce from Lemma 3 in [20] that if we put

$$H_n^-(x) := \mathbf{P}\left(\sum_{k=1}^{n} e^{S_k} \le x \,\middle|\, \widetilde{M}_n < 0\right), \qquad H_n^+(x) := \mathbf{P}\left(\sum_{k=0}^{n} e^{-S_k} \le x \,\middle|\, L_n \ge 0\right),$$

and

$$H^-(x) := \mathbf{P}^-(\eta_1 < x), \qquad H^+(x) := \mathbf{P}^+(\eta_2 < x),$$

then under the condition (19)

$$H_n^{\pm}(x) \Rightarrow H^{\pm}(x) \quad \text{ as } \ n\to\infty, \tag{27}$$

where the symbol $\Rightarrow$ denotes convergence at all continuity points of the limiting function.

In what follows we will often use the following statement (see e.g. Lemma 2.1 in [1], Theorem 8.9.12 in [7], and Lemma 2 in [20]).

Let

$$\lambda_n(x) := \mathbf{P}\left(L_n \ge -x\right), \qquad \widetilde{\mu}_n(x) := \mathbf{P}\left(\widetilde{M}_n < x\right), \qquad x\ge 0.$$

Lemma 1. *Under Sptizer's condition* (19) *there exist slowly varying at infinity functions $l_1(n)$ and $l_2(n)$, related by $l_1(n)l_2(n)\sim\pi^{-1}\sin\pi\rho$, $n\to\infty$, such that*

$$\mathbf{P}\left(L_n\ge 0\right)\sim n^{\rho-1}l_1(n), \quad \mathbf{P}\left(\widetilde{M}_n<0\right)\sim n^{-\rho}l_2(n) \quad \text{as } \ n\to\infty. \tag{28}$$

Moreover, there are absolute constants $C_1>0$, $C_2>0$ such that for all $n\ge1$ and $x\ge0$

$$\lambda_n(x)\le C_1V(x)\,\mathbf{P}\left(L_n\ge0\right), \qquad \widetilde{\mu}_n(x)\le C_2U(x)\,\mathbf{P}\left(\widetilde{M}_n<0\right). \tag{29}$$

In (28) and in the rest of the paper, notation $a_n\sim b_n$ means that $a_n/b_n\to1$ as $n\to\infty$.

Let $\{S_n^-\}_{n\ge0}$ and $\{S_n^+\}_{n\ge0}$ be two independent copies of $\{S_n\}_{n\ge0}$, and let

$$L_n^+ := \min_{0\le r\le n} S_r^+, \qquad \widetilde{M}_n^- := \max_{1\le l\le n} S_l^-.$$

Introduce the probability distributions

$$\mathbf{P}_{-,+} := \mathbf{P}^-\times\mathbf{P}^+, \quad \mathbf{P}_{\cdot,+} := \mathbf{P}\times\mathbf{P}^+, \quad \mathbf{P}_{-,\cdot} := \mathbf{P}^-\times\mathbf{P}$$

on the sample space $\mathbb{R}^\infty \times \mathbb{R}^\infty$ of the pair $\big(\{S_n^-\}_{n\ge 0},\, \{S_n^+\}_{n\ge 0}\big)$, where $\mathbf{P}$ is the distribution of the original sequence $\{S_n\}_{n\ge 0}$ and the measures $\mathbf{P}^\pm$ are specified by (24), (25), and let $\mathbf{E}_{-,+}$, $\mathbf{E}_{\cdot,+}$, and $\mathbf{E}_{-,\cdot}$ be the expectation operators under the respective measures.

We will call an array of r.v.'s $\{G_{l,r};\, l,r\in\mathbb{N}\}$ *adapted* if, for any pair of indices $l,r\in\mathbb{N}$, the r.v. $G_{l,r}$ is measurable with respect to the σ-algebra $\sigma(S_1^-,\dots,S_l^-)\otimes\sigma(S_1^+,\dots,S_r^+)$. The following result is contained in Lemma 3 in [20].

Lemma 2. *Let Spitzer's condition* (19) *hold, and let* $\{G_{l,r};\, l,r\in\mathbb{N}\}$ *be an adapted array of uniformly bounded r.v.'s. If the following limit exists:*

$$\lim_{l,r\to\infty} G_{l,r} =: G \qquad \mathbf{P}_{-,+}\text{-}a.s.,$$

then

$$\lim_{l,r\to\infty}\mathbf{E}\left[G_{l,r}\,|\,\widetilde{M}_l^- < 0,\, L_r^+\ge 0\right] = \mathbf{E}_{-,+}G. \tag{30}$$

The next statement is a slight modification of Lemma 2.5 in [1] and can be proved by the same arguments as used there.

Lemma 3. *Let Spitzer's condition* (19) *hold, and let* $\{G_{l,r},\, l,r\in\mathbb{N}\}$ *be an adapted array of uniformly bounded r.v.'s. If the following limit exists:*

$$\lim_{r\to\infty} G_{l,r} I\{\widetilde{M}_l^- <0\} =: G_l^+ I\{\widetilde{M}_l^- <0\} \qquad \mathbf{P}_{\cdot,+}\text{-}a.s.,$$

then

$$\lim_{r\to\infty}\mathbf{E}\left[G_{l,r} I\{\widetilde{M}_l^- <0\}|\, L_r^+\ge 0\right] = \mathbf{E}_{\cdot,+} G_l^+ I\{\widetilde{M}_l^- <0\},$$

and if

$$\lim_{l\to\infty} G_{l,r} I\{L_r^+\ge 0\} =: G_r^- I\{L_r^+\ge 0\} \qquad \mathbf{P}_{-,\cdot}\text{-}a.s.,$$

then

$$\lim_{l\to\infty}\mathbf{E}\left[G_{l,r} I\{L_r^+\ge 0\}|\, \widetilde{M}_l^- <0\right] = \mathbf{E}_{-,\cdot} G_r^- I\{L_r^+\ge 0\}.$$

The following result was proved in Lemma 2.2 of [1]. Denote by

$$\tau(n) := \min\big\{k\ge 0:\ S_k\le S_l,\ l\in[0,n]\big\}$$

the left-most point at which the random walk $\{S_n\}$ attains its minimum value on the time interval $[0,n]$.

Lemma 4. *Let Spitzer's condition* (19) *hold, and let* $u(x)\ge 0$, $x\ge 0$, *be a nonincreasing function such that* $\int_0^\infty u(x)dx<\infty$. *Then, for every* $\varepsilon>0$, *there exists an integer* J *such that for all* $n\ge J$

$$\sum_{p=J}^{n}\mathbf{E}\left[u(-S_p);\, \tau(p)=p\right]\mathbf{P}\left(L_{n-p}\ge 0\right)\le \varepsilon\mathbf{P}\left(L_n\ge 0\right).$$

Introduce the r.v.'s

$$G_r^+(j) := \frac{e^{-S_{j-r}^+} I\{j \geq r\} + e^{S_{r-j}^-} I\{j < r\}}{\sum_{p=1}^r e^{S_p^-} + \eta_2^+} \tag{31}$$

and

$$G_r^-(j) := \frac{e^{S_{j-r}^-} I\{j > r\} + e^{-S_{r-j}^+} I\{j \leq r\}}{\eta_1^- + \sum_{p=0}^r e^{-S_p^+}},$$

where η_1^- and η_2^+ are defined as in (26), but for the random walks $\{S_n^-\}_{n\geq 0}$ and $\{S_n^+\}_{n\geq 0}$, respectively. Note that $0 < G_r^\pm \leq 1$ and, in view of (26), $G_r^+(j)$ and $G_r^-(j)$ are a.s. positive under the measures $\mathbf{P}_{\cdot,+}$ and $\mathbf{P}_{-,\cdot}$, respectively. Set

$$\widetilde{L}_n^+ := \min_{1\leq p\leq n} S_p^+ \tag{32}$$

and put

$$c_j := \sum_{l=0}^\infty \mathbf{E}_{\cdot,+} G_l^+(j) I\{\widetilde{M}_l^- < 0\}, \qquad d_j := \sum_{q=1}^j \sum_{r=0}^\infty \mathbf{E}_{-,\cdot} G_r^-(q) I\{\widetilde{L}_r^+ > 0\}.$$

One can easily verify that c_j and d_j are finite for any $j = 0, 1, \ldots$. Thus,

$$\begin{aligned} c_j &\leq j + 1 + \sum_{l=j+1}^\infty \mathbf{E}_{\cdot,+} e^{S_{l-j}^-} I\{\widetilde{M}_l^- < 0\} = j + 1 + \sum_{p=1}^\infty \mathbf{E}\, e^{S_p} I\{\widetilde{M}_p < 0\} \\ &= j + 1 + \sum_{p=1}^\infty \mathbf{E}\, e^{S_p} I\{S_1 < 0, \ldots, S_p < 0\} < \infty \end{aligned}$$

(see Section 17, D2 in [19]).

Now we are ready to formulate and prove the following statement.

Theorem 4. *Let Spitzer's condition* (19) *hold. Then for any fixed* $j \geq 0$

$$\lim_{n\to\infty} \frac{\mathbf{E}\, N_n(j)}{n\mathbf{P}\,(L_n \geq 0)} = \frac{c_j}{\rho} \tag{33}$$

and

$$\lim_{n\to\infty} \frac{\mathbf{E}\, N_n(n-j)}{\mathbf{P}\,(\widetilde{M}_n < 0)} = d_j. \tag{34}$$

Remark 5. In view of (28), the relations (33) and (34) can be rewritten as

$$\mathbf{E}\, N_n(j) \sim c_j \rho^{-1} n^\rho l_1(n), \quad \mathbf{E}\, N_n(n-j) \sim d_j n^{-\rho} l_2(n) \quad \text{as } n \to \infty.$$

Proof. To prove Theorem 4, we have to evaluate the sum (18) of expectations of the form

$$\mathbf{E}\, e^{-S_j} W_k^{-1} = \sum_{l=0}^k \mathbf{E}\, e^{-S_j} W_k^{-1} I\{\tau(k) = l\}. \tag{35}$$

The key idea both in this proof and also in that of Theorem 5 is quite similar to that of the Laplace method: the main contribution to the expectation (35) comes from the event where j is close to $\tau(k)$ (for other values of $j \le k$, the quantity e^{-S_j} will typically be quite small compared to W_k).

First we will show that, for each fixed $\varepsilon > 0$, there exists a $J = J(\varepsilon)$ such that for all $j \ge 0$ and all $k \ge J + j$

$$\mathbf{E}\, e^{-S_j} W_k^{-1} I\{\tau(k) \ge J + j\} \le \varepsilon \mathbf{P}\,(L_{k-j} \ge 0). \tag{36}$$

Indeed, as $W_k \ge e^{-S_{\tau(k)}}$, we have

$$\begin{aligned} \mathbf{E}\, e^{-S_j} W_k^{-1} I\{\tau(k) \ge J + j\} &\le \mathbf{E}\, e^{S_{\tau(k)} - S_j} I\{\tau(k) \ge J + j\} \\ &= \sum_{p=J}^{k-j} \mathbf{E}\, e^{S_{p+j} - S_j} I\{\tau(k) = p + j\} \le \sum_{p=J}^{k-j} \mathbf{E}\, e^{S_p} I\{\tau(k-j) = p\} \\ &= \sum_{p=J}^{k-j} \mathbf{E}\left[e^{S_p} I\{\tau(p) = p\}\right] \mathbf{P}\,(L_{k-j-p} \ge 0), \end{aligned}$$

and to get the desired statement it remains to apply Lemma 4 with $u(x) = e^{-x}$.

The next step is to demonstrate that for any fixed $j \ge 0$, $l \ge 1$

$$\lim_{k\to\infty} \frac{\mathbf{E}\, e^{-S_j} W_k^{-1} I\{\tau(k) = l\}}{\mathbf{P}\,(L_k \ge 0)} = \mathbf{E}_{\cdot,+} G_l^+(j) I\{\widetilde{M}_l^- < 0\}. \tag{37}$$

But this is an easy consequence of Lemma 3. Indeed, assume first that $j \ge l$. Then for the r.v.'s $G_{l,r}(j)$ defined for $r \ge j - l$ by

$$G_{l,k-l}(j) := \frac{e^{-S_{j-l}^+}}{\sum_{p=1}^{l} e^{S_p^-} + \sum_{q=0}^{k-l} e^{-S_q^+}} \le 1, \qquad k \ge j \tag{38}$$

(for $r < j - l$ one can put $G_{l,r}(j) \equiv 1$), we have

$$\begin{aligned} \mathbf{E}\, e^{-S_j} W_k^{-1} I\{\tau(k) = l\} &= \mathbf{E}\, \frac{e^{S_{\tau(k)} - S_j}}{\sum_{p=0}^{k} e^{S_{\tau(k)} - S_p}} I\{\tau(k) = l\} \\ &= \mathbf{E}\, G_{l,k-l}(j) I\{\widetilde{M}_l^- < 0,\, L_{k-l}^+ \ge 0\} \\ &= \mathbf{E}\left[G_{l,k-l}(j) I\{\widetilde{M}_l^- < 0\} \middle| L_{k-l}^+ \ge 0\right] \mathbf{P}\,(L_{k-l} \ge 0) \end{aligned}$$

(here the second relation follows from the duality principle: we use the "time-reversed random walk" on $[0, l]$).

It is evident that, as $k \to \infty$,

$$G_{l,k-l}(j) I\{\widetilde{M}_l^- < 0\} \to G_l^+(j) I\{\widetilde{M}_l^- < 0\} \qquad \mathbf{P}_{\cdot,+}\text{-a.s.},$$

and therefore by Lemma 3

$$\lim_{k\to\infty} \mathbf{E}\left[G_{l,k-l}(j) I\{\widetilde{M}_l^- < 0\} \middle| L_{k-l}^+ \ge 0\right] = \mathbf{E}_{\cdot,+} G_l^+(j) I\{\widetilde{M}_l^- < 0\}. \tag{39}$$

On the other hand, in view of (28) for each fixed l

$$\lim_{k\to\infty} \frac{\mathbf{P}\,(L_{k-l} \geq 0)}{\mathbf{P}\,(L_k \geq 0)} = 1. \tag{40}$$

Combining this with (39) gives (37). The case $j < l$ can be treated in a similar way.

Now everything is ready to complete the proof of the first part of the theorem. It follows from (35), (36) and (37) that, for each fixed $j \geq 0$,

$$\mathbf{E}\, e^{-S_j} W_k^{-1} \sim c_j \mathbf{P}\,(L_k \geq 0) \quad \text{as} \;\; k \to \infty. \tag{41}$$

Therefore, for a fixed $\varepsilon > 0$ there exists a $K(\varepsilon) < \infty$ such that for all $K \geq K(\varepsilon)$ and $n > K$

$$\begin{aligned}(1-\varepsilon)c_j \sum_{k=K+1}^{n-1} \mathbf{P}\,(L_k \geq 0) \leq \mathbf{E}\, N_n(j) &= \sum_{k=j+1}^{K} \mathbf{E}\, e^{-S_j} W_k^{-1} + \sum_{k=K+1}^{n-1} \mathbf{E}\, e^{-S_j} W_k^{-1} \\ &\leq (K-j) + (1+\varepsilon)c_j \sum_{k=K+1}^{n-1} \mathbf{P}\,(L_k \geq 0).\end{aligned} \tag{42}$$

By (28) and Karamata's theorem (see e.g. Section 1.6 in [7])

$$\sum_{k=K+1}^{n-1} \mathbf{P}\,(L_k \geq 0) \sim \frac{n}{\rho} \mathbf{P}\,(L_n \geq 0) \quad \text{as} \;\; n \to \infty. \tag{43}$$

This together with (42) completes the proof of (33).

Now we will prove (34). Let $\{S_n^*\}_{n\geq 0} \overset{d}{=} \{-S_n\}_{n\geq 0}$ be the "reflected" random walk. By the duality principle, for each fixed $q \leq j$

$$\begin{aligned}\mathbf{E}\, e^{-S_{n-j}} W_{n-q}^{-1} &= \mathbf{E}\, \frac{e^{-S_{n-j}}}{\sum_{p=0}^{n-q} e^{-S_{n-q-p}}} \\ &= \mathbf{E}\, \frac{e^{S_{n-q}-S_{n-j}}}{\sum_{p=0}^{n-q} e^{S_{n-q}-S_{n-q-p}}} = \mathbf{E}\, \frac{e^{-S_{j-q}^*}}{\sum_{p=0}^{n-q} e^{-S_p^*}} = \mathbf{E}\, e^{-S_{j-q}^*} (W_{n-q}^*)^{-1}\end{aligned} \tag{44}$$

(with an obvious definition of W_{n-q}^*).

Next we set

$$L_n^* := \min_{0\leq k\leq n} S_k^*, \qquad \widetilde{M}_n^* := \max_{1\leq k\leq n} S_k^*$$

and observe that, as $n \to \infty$,

$$\mathbf{P}\,(L_n^* \geq 0) = \mathbf{P}\,(M_n \leq 0) \sim e^{\phi} \mathbf{P}\,(\widetilde{M}_n < 0). \tag{45}$$

Indeed, putting

$$\chi := \inf\{k \geq 1 : \; S_k \geq 0\}, \qquad \widetilde{\chi} := \inf\{k \geq 1 : \; S_k > 0\},$$

we get from the factorization identities that for $|z|<1$

$$1-\mathbf{E}\,z^{\widetilde{\chi}}=\exp\biggl\{\sum_{n=0}^{\infty}\frac{z^n}{n}\,\mathbf{P}\,(S_n>0)\biggr\},\qquad 1-\mathbf{E}\,z^{\chi}=\exp\biggl\{\sum_{n=0}^{\infty}\frac{z^n}{n}\,\mathbf{P}\,(S_n\ge 0)\biggr\}$$

(see e.g. Corollary 4, § 16 in [9]). Dividing both sides of these identities by $1-z=e^{\ln(1-z)}$, we obtain

$$\begin{aligned}\sum_{n=0}^{\infty}z^n\mathbf{P}\,(M_n\le 0)&=\sum_{n=0}^{\infty}z^n\mathbf{P}\,(\widetilde{\chi}>n)=\frac{1-\mathbf{E}\,z^{\widetilde{\chi}}}{1-z}\\&=\exp\biggl\{-\sum_{n=1}^{\infty}\frac{z^n}{n}\,\mathbf{P}\,(S_n>0)+\sum_{n=1}^{\infty}\frac{z^n}{n}\biggr\}=\exp\biggl\{\sum_{n=1}^{\infty}\frac{z^n}{n}\,\mathbf{P}\,(S_n\le 0)\biggr\},\end{aligned}$$

and similarly

$$\sum_{n=0}^{\infty}z^n\mathbf{P}\,(\widetilde{M}_n<0)=\sum_{n=0}^{\infty}z^n\mathbf{P}\,(\chi>n)=\exp\biggl\{\sum_{n=1}^{\infty}\frac{z^n}{n}\,\mathbf{P}\,(S_n<0)\biggr\}.$$

Therefore

$$\sum_{n=0}^{\infty}z^n\mathbf{P}\,(M_n\le 0)=e^{\phi(z)}\sum_{n=0}^{\infty}z^n\mathbf{P}\,(\widetilde{M}_n<0),\qquad \phi(z):=\sum_{n=1}^{\infty}\frac{z^n}{n}\,\mathbf{P}\,(S_n=0).$$

To get (45), it remains to use (28) and Karamata's Tauberian theorem (see e.g. Corollary 1.7.3 in [7]), noting that $\phi(z)\to\phi$ as $z\nearrow 1$.

Now from (41) and (45) we obtain that, as $n\to\infty$,

$$\mathbf{E}\,e^{-S^*_{j-q}}(W^*_{n-q})^{-1}\sim c^*_{j-q}\mathbf{P}\,\bigl(L^*_{n-q}\ge 0\bigr)\sim c^*_{j-q}e^{\phi}\mathbf{P}\,(\widetilde{M}_n<0),$$

where, with a natural definition of $\mathbf{E}^*_{\cdot,+}$ and with $\widetilde{L}^+_r$ defined in (32), due to the definitions (24) and (25), one has

$$e^{\phi}c^*_{j-q}=e^{\phi}\sum_{l=0}^{\infty}\mathbf{E}^*_{\cdot,+}G^{*+}_l(j-q)I\{\widetilde{M}^{*-}_l<0\}=\sum_{r=0}^{\infty}\mathbf{E}_{-,\cdot}G^-_r(j-q)I\{\widetilde{L}^+_r>0\}.$$

Therefore we have from (18) and (44) that, as $n\to\infty$,

$$\begin{aligned}\mathbf{E}\,N_n(n-j)&=\sum_{k=n-j}^{n-1}\mathbf{E}\,e^{-S_{n-j}}W_k^{-1}=\sum_{q=1}^{j}\mathbf{E}\,e^{-S^*_{j-q}}(W^*_{n-q})^{-1}\sim\mathbf{P}\,(L^*_n\ge 0)\sum_{q=1}^{j}c^*_{j-q}\\&\sim\mathbf{P}\,(L^*_n\ge 0)e^{-\phi}\sum_{q=1}^{j}\sum_{r=0}^{\infty}\mathbf{E}_{-,\cdot}G^-_r(j-q)I\{\widetilde{L}^+_r>0\}\sim d_j\mathbf{P}\,(\widetilde{M}_n<0),\end{aligned}$$

as desired. Theorem 4 is proved. □

The next theorem describes the asymptotic behavior of the expectation $\mathbf{E}\, N_n(j)$ when $\min\{j, n-j\} \to \infty$.

Theorem 5. *Let Spitzer's condition* (19) *be satisfied. Then*

$$\lim_{j,n-j\to\infty} \frac{\mathbf{E}\, N_n(j)}{(n-j)\mathbf{P}\,(\widetilde{M}_j < 0)\mathbf{P}\,(L_{n-j} \ge 0)} = \frac{1}{\rho}. \tag{46}$$

Remark 6. In view of (28), the assertion of the theorem can be rewritten as

$$\mathbf{E}\, N_n(j) \sim \rho^{-1} j^{-\rho} l_2(j)(n-j)^{\rho} l_1(n-j) \quad \text{as} \;\; j, n-j \to \infty.$$

It follows that, for any fixed $\varepsilon \in (0, 1/2)$, we have for $t \in [\varepsilon, 1-\varepsilon]$

$$\mathbf{E}\, N_n(\lfloor nt \rfloor) \sim \frac{\sin \pi\rho}{\pi\rho} \left(\frac{1-t}{t} \right)^{\rho} \quad \text{as} \;\; n \to \infty.$$

It is interesting to compare this with the corresponding (obvious) asymptotics for the case when $w(j) \equiv 1$: then $\mathbf{E}\, N_n(\lfloor nt \rfloor) \sim -\ln t$ (of course, the functions of t on the right-hand sides of the both relations are densities on $(0, 1)$).

In the case when $\mathbf{E}\,\theta = 0$, $\mathbf{E}\,\theta^2 < \infty$, we don't even need to bound the value j/n away from 0 and 1: in that case, from the asymptotic behaviour of the denominators in (46) (see e.g. p.94 in [9]), we get

$$\mathbf{E}\, N_n(j) \sim \frac{2}{\pi} \left(\frac{n-j}{j} \right)^{1/2} \quad \text{as} \;\; j, n-j \to \infty.$$

Note also that the assertions (33), (34) of Theorem 4 can be viewed as the "boundary cases" of (46): there is a "smooth transition" between these asymptotics.

We split the proof of the theorem into several steps. As we said before, the main contribution to the expectation $\mathbf{E}\, e^{-S_j} W_k^{-1}$ from the sum (18) comes from the event where j is close to $\tau(k)$. So first we will show that the contribution from the complementary event is negligibly small indeed.

Lemma 5. *Under Spitzer's condition* (19)*, for any $\varepsilon > 0$ there exists a $J = J(\varepsilon) < \infty$ such that for all $j \ge J$ and $k - j \ge J$*

$$\mathbf{E}\left[e^{S_{\tau(k)} - S_j};\, |\tau(k) - j| \ge J \right] \le \varepsilon \mathbf{P}\,(\widetilde{M}_j < 0)\mathbf{P}\,(L_{k-j} \ge 0). \tag{47}$$

Proof. Fix a $J > 0$ and choose a $j \ge J$ and a $k \ge j + J$. We have

$$\mathbf{E}\left[e^{S_{\tau(k)} - S_j};\, |\tau(k) - j| \ge J \right] = R_1 + R_2,$$

where

$$R_1 := \sum_{t=0}^{j-J} \mathbf{E}\left[e^{S_{\tau(k)} - S_j};\, \tau(k) = t \right], \qquad R_2 := \sum_{t=j+J}^{k} \mathbf{E}\left[e^{S_{\tau(k)} - S_j};\, \tau(k) = t \right].$$

First consider R_2. For $t \geq j$ we get

$$\begin{aligned}\mathbf{E}\left[e^{S_{\tau(k)}-S_j};\,\tau(k)=t\right] &= \mathbf{E}\left[e^{S_t-S_j};\,\min_{0\le p\le t-1} S_p > S_t,\,\min_{t\le p\le k} S_p \ge S_t\right]\\ &= \mathbf{E}\left[e^{S_t-S_j};\,\min_{0\le p\le t-1} S_p > S_t\right]\mathbf{P}\left(L_{k-t}\ge 0\right)\\ &= \mathbf{E}\left[e^{S_{t-j}};\,\max_{1\le p\le t} S_p < 0\right]\mathbf{P}\left(L_{k-t}\ge 0\right)\end{aligned}$$

by the duality principle. Defining for each $l \geq 0$ the shifted random walk

$$\{S_p^{(l)} := S_{l+p} - S_l\}_{p\ge 0},$$

we obtain from (29) that

$$\begin{aligned}\mathbf{E}\left[e^{S_{t-j}};\,\max_{1\le p\le t} S_p < 0\right] &= \mathbf{E}\left[e^{S_{t-j}}\mathbf{P}\left(\max_{1\le p\le j} S_p^{(t-j)} < -S_{t-j}\,\Big|\, S_{t-j}\right);\,\max_{1\le p\le t-j} S_p < 0\right]\\ &= \mathbf{E}\left[e^{S_{t-j}}\widetilde{\mu}_j(-S_{t-j});\,\widetilde{M}_{t-j}<0\right]\\ &\le C_2\,\mathbf{P}\left(\widetilde{M}_j<0\right)\mathbf{E}\left[e^{S_{t-j}}U(-S_{t-j});\,\widetilde{M}_{t-j}<0\right].\end{aligned}$$

Hence

$$\begin{aligned}R_2 &\le C_2\,\mathbf{P}\left(\widetilde{M}_j<0\right)\sum_{t=j+J}^{k}\mathbf{E}\left[e^{S_{t-j}}U(-S_{t-j});\,\widetilde{M}_{t-j}<0\right]\mathbf{P}\left(L_{k-t}\ge 0\right)\\ &= C_2\,\mathbf{P}\left(\widetilde{M}_j<0\right)\sum_{p=J}^{k-j}\mathbf{E}\left[e^{S_p}U(-S_p);\,\widetilde{M}_p<0\right]\mathbf{P}\left(L_{k-j-p}\ge 0\right).\end{aligned}$$

Since $U(x)$ is a renewal function, we have $U(x) = O(x)$, $x \to \infty$. Thus, there exists a constant C_3 such that $e^{-x}U(x) \le u(x) := C_3 e^{-x/2}$ for all $x > 0$. Since $\int_0^\infty u(x)\,dx < \infty$, it follows from Lemma 4 and the duality principle that, for every $\varepsilon > 0$, there exists a $J_1 = J_1(\varepsilon) < \infty$ such that for all $k - j > J_1$

$$\sum_{p=J_1}^{k-j}\mathbf{E}\left[e^{S_p}U(-S_p);\,\widetilde{M}_p<0\right]\mathbf{P}\left(L_{k-j-p}\ge 0\right) \le \frac{\varepsilon}{2C_2}\mathbf{P}\left(L_{k-j}\ge 0\right).$$

Thus, for $k - j > J \ge J_1$,

$$R_2 \le \frac{\varepsilon}{2}\,\mathbf{P}\left(\widetilde{M}_j<0\right)\mathbf{P}\left(L_{k-j}\ge 0\right). \tag{48}$$

Now we will evaluate R_1. For $t < j$ we get

$$\begin{aligned}\mathbf{E}\left[e^{S_{\tau(k)}-S_j};\,\tau(k)=t\right] &= \mathbf{E}\left[e^{S_t-S_j};\,\min_{0\le p\le t-1} S_p > S_t;\,\min_{t\le p\le k} S_p \ge S_t\right]\\ &= \mathbf{E}\left[e^{S_t-S_j};\,\min_{t\le p\le k} S_p \ge S_t\right]\mathbf{P}\left(\widetilde{M}_t<0\right)\\ &= \mathbf{E}\left[e^{-S_{j-t}};\,\min_{0\le p\le k-t} S_p \ge 0\right]\mathbf{P}\left(\widetilde{M}_t<0\right),\end{aligned}$$

where to obtain the second relation we again used the duality principle. Arguing as before, we see that

$$\begin{aligned}
&\mathbf{E}\left[e^{-S_{j-t}};\ \min_{0\le p\le k-t} S_p \ge 0\right] \\
&\qquad = \mathbf{E}\left[e^{-S_{j-t}}\mathbf{P}\left(\min_{0\le p\le k-j} S_p^{(j-t)} \ge -S_{j-t} \,\Big|\, S_{j-t}\right);\ \min_{0\le p\le j-t} S_p \ge 0\right] \\
&\qquad = \mathbf{E}\left[e^{-S_{j-t}}\lambda_{k-j}(S_{j-t});\ L_{j-t} \ge 0\right] \\
&\qquad \le C_1\,\mathbf{P}\left(L_{k-j} \ge 0\right)\mathbf{E}\left[e^{-S_{j-t}}V(S_{j-t});\ L_{j-t} \ge 0\right].
\end{aligned}$$

Hence

$$\begin{aligned}
R_1 &\le C_1\,\mathbf{P}\left(L_{k-j} \ge 0\right)\sum_{t=0}^{j-J}\mathbf{E}\left[e^{-S_{j-t}}V(S_{j-t});\ L_{j-t} \ge 0\right]\mathbf{P}\left(\widetilde{M}_t < 0\right) \\
&\qquad = C_1\,\mathbf{P}\left(L_{k-j} \ge 0\right)\sum_{p=J}^{j}\mathbf{E}\left[e^{-S_p}V(S_p);\ L_p \ge 0\right]\mathbf{P}\left(\widetilde{M}_{j-p} < 0\right).
\end{aligned}$$

From this bound one can deduce, using Lemma 4 and the same argument as the one employed to evaluate R_2, that for every $\varepsilon > 0$ there exists a $J_2(\varepsilon) < \infty$ such that for all $j > J \ge J_2$

$$R_1 \le \frac{\varepsilon}{2}\,\mathbf{P}\,(\widetilde{M}_j < 0)\mathbf{P}\,(L_{k-j} \ge 0). \tag{49}$$

Combining (48) with (49) and setting $J := \max\{J_1, J_2\}$ completes the proof of Lemma 5. □

Next we evaluate the contributions to the expectations of interest from the events where $\tau(k)$ is equal to a fixed number close to j.

Lemma 6. *Under Sptizer's condition* (19), *for any fixed* $r \in \mathbb{Z}$

$$\lim_{j,k-j\to\infty}\frac{\mathbf{E}\left[e^{-S_j}W_k^{-1};\ \tau(k) = j + r\right]}{\mathbf{P}\,(\widetilde{M}_j < 0)\mathbf{P}\,(L_{k-j} \ge 0)} = \mathbf{E}_{-,+}\frac{e^{S_r^-}I\{r \ge 0\} + e^{-S_{-r}^+}I\{r < 0\}}{\eta_1^- + \eta_2^+}, \tag{50}$$

where η_1^- *and* η_2^+ *are independent r.v.'s defined as in* (26), *but for the independent random walks* $\{S_n^-\}_{n\ge0}$ *and* $\{S_n^+\}_{n\ge0}$, *respectively.*

Proof. For $0 \le r \le k - j$ put

$$G_{j+r,k-j-r} := \frac{e^{S_r^-}}{\sum_{p=1}^{j+r} e^{S_p^-} + \sum_{p=0}^{k-j-r} e^{-S_p^+}}.$$

Then

$$\begin{aligned}
&\mathbf{E}\left[e^{-S_j}W_k^{-1};\ \tau(k) = j + r\right] \\
&\qquad = \mathbf{E}\left[\frac{e^{S_{j+r}-S_j}}{\sum_{p=0}^{k} e^{S_{j+r}-S_p}};\ \min_{0\le p\le j+r-1} S_p > S_{j+r};\ \min_{j+r\le p\le k} S_p \ge S_{j+r}\right] \\
&\qquad = \mathbf{E}\left[G_{j+r,k-j-r};\ \widetilde{M}_{j+r}^- < 0,\ L_{k-j-r}^+ \ge 0\right] \\
&\qquad = \mathbf{E}\left[G_{j+r,k-j-r} \,\big|\, \widetilde{M}_{j+r}^- < 0,\ L_{k-j-r}^+ \ge 0\right]\mathbf{P}\,(\widetilde{M}_{j+r} < 0)\mathbf{P}\,(L_{k-j-r} \ge 0).
\end{aligned}$$

Clearly, $0 < G_{j+r,k-j-r} \leq 1$ and

$$\lim_{j,k-j\to\infty} G_{j+r,k-j-r} = \frac{e^{S_r^-}}{\eta_1^- + \eta_2^+} \qquad \mathbf{P}_{-,+}\text{-a.s.}$$

Hence, applying Lemma 2 and recalling (28) and the properties of regularly varying functions (cf. (40)), we get (50) for $r \geq 0$. The proof of (50) in the case $r < 0$ is almost identical. Lemma 6 is proved. $\square$

Proof of Theorem 5. For a fixed $\varepsilon > 0$ let $J = J(\varepsilon)$ be such that (47) holds true. For $j \geq J$ and $n - j \geq J + 1$ we have from (18) that

$$\mathbf{E}\, N_n(j) = R_3 + R_4 + R_5,$$

where

$$R_3 := \sum_{k=j}^{j+J-1} \mathbf{E}\, e^{-S_j} W_k^{-1}, \qquad R_4 := \sum_{k=j+J}^{n-1} \mathbf{E}\left[e^{-S_j} W_k^{-1};\, |\tau(k) - j| < J\right]$$

and

$$R_5 := \sum_{k=j+J}^{n-1} \mathbf{E}\left[e^{-S_j} W_k^{-1};\, |\tau(k) - j| \geq J\right].$$

We evaluate the quantities R_i, $i = 3, 4, 5$, separately. First observe that, in view of (34) (with n replaced by k), there exists a constant C_3 such that for all sufficiently large j

$$R_3 \leq C_3 J\, \mathbf{P}\,(\widetilde{M}_j < 0),$$

and since

$$(n-j)\, \mathbf{P}\,(L_{n-j} \geq 0) \sim (n-j)^{\rho} l_1(n-j) \to \infty \quad \text{ as } \; n - j \to \infty,$$

it follows that

$$R_3 = o\Big((n-j)\, \mathbf{P}\,(\widetilde{M}_j < 0)\mathbf{P}\,(L_{n-j} \geq 0)\Big) \quad \text{ as } \; n - j \to \infty. \tag{51}$$

Further, using the obvious inequality $W_k \geq e^{-S_{\tau(k)}}$ and the bound (47) together with (28) and Karamata's theorem, we get for $j \geq J$ and some positive absolute constant C_5 that

$$\begin{aligned} R_5 &\leq \varepsilon \mathbf{P}\,(\widetilde{M}_j < 0) \sum_{k=j+J}^{n-1} \mathbf{P}\,(L_{k-j} \geq 0) \\ &= \varepsilon \mathbf{P}\,(\widetilde{M}_j < 0) \sum_{p=J}^{n-j-1} \mathbf{P}\,(L_p \geq 0) \leq \varepsilon C_5 (n-j)\, \mathbf{P}\,(\widetilde{M}_j < 0)\mathbf{P}\,(L_{n-j} \geq 0), \end{aligned}$$

and therefore

$$\frac{R_5}{(n-j)\, \mathbf{P}\,(\widetilde{M}_j < 0)\mathbf{P}\,(L_{k-j} \geq 0)} \leq \varepsilon C_5.$$

Finally, set

$$E_J := \mathbf{E}_{-,+} \frac{1 + \sum_{r=1}^{J-1} \left(e^{S_r^-} + e^{-S_r^+}\right)}{\eta_1^- + \eta_2^+} .$$

Using Lemma 6, the relation (28) and the properties of regularly varying functions, we see that, as $\min\{j, n-j\} \to \infty$,

$$\begin{aligned} R_4 &\sim E_J \mathbf{P}\left(\widetilde{M}_j < 0\right) \sum_{k=j+J}^{n-1} \mathbf{P}\left(L_{k-j} \geq 0\right) \\ &\sim E_J \mathbf{P}\left(\widetilde{M}_j < 0\right) \sum_{p=J}^{n-j-1} \mathbf{P}\left(L_p \geq 0\right) \sim E_J \mathbf{P}\left(\widetilde{M}_j < 0\right) \rho^{-1}(n-j)\, \mathbf{P}\left(L_{n-j} \geq 0\right). \end{aligned}$$

Since $\lim_{J \to \infty} E_J = 1$ by the dominated convergence theorem, the assertion of Theorem 5 immediately follows from the above relation for R_4 and the bounds for R_3 and R_5. $\square$

3.2 The asymptotic behavior of the distribution of $\mathbf{E}_w N_n(j)$

Unfortunately, our description of the asymptotic behavior of $\mathbf{E}_w N_n(j)$ will be less detailed than that of $\mathbf{E}\, N_n(j)$. We will be able to describe the distribution of the r.v. $\mathbf{E}_w N_n(j)$ only for j located either to the right or in a small left vicinity of the random epoch $\tau(n)$.

Theorem 6. *Let Spitzer's condition* (19) *be satisfied and $j = j(n)$ be an arbitrary (random) sequence with the property that $(\tau(n) - j)_+ = o(n)$ in probability as $n \to \infty$. Then*

$$\mathbf{P}\left(\frac{e^{S_j - S_{\tau(n)}}}{n-j} \mathbf{E}_w N_n(j) < x\right) \Rightarrow \mathbf{P}_{-,+}\left(\frac{1}{\eta_1^- + \eta_2^+} < x\right), \tag{52}$$

where η_1^- and η_2^+ are r.v.'s defined as in (26), *but for the independent random walks $\{S_n^-\}_{n \geq 0}$ and $\{S_n^+\}_{n \geq 0}$, respectively.*

Proof. Since the r.v.'s W_n (see (5)) are increasing in n, we have from (17) the following lower bound:

$$\mathbf{E}_w N_n(j) \geq (n-j)\, e^{-S_j} W_n^{-1} = \frac{(n-j)\, e^{S_{\tau(n)} - S_j}}{\sum_{k=0}^{n} e^{S_{\tau(n)} - S_k}} .$$

Now we will derive an upper bound for $\mathbf{E}_w N_n(j)$. To this end observe that, according to (26), for any fixed $\varepsilon > 0$ and $\delta > 0$ there exists a $J < \infty$ such that

$$\mathbf{P}^+\left(\sum_{k=J}^{\infty} e^{-S_k} > \delta\right) \leq \varepsilon. \tag{53}$$

Clearly, for any $j \in [\tau(n), n-1]$

$$\begin{aligned}\mathbf{E}_w N_n(j) &\le e^{S_{\tau(n)}-S_j}(\tau(n)+J-j)_+ + e^{-S_j}(n-j)W^{-1}_{\tau(n)+J} \\ &= e^{S_{\tau(n)}-S_j}\left[(\tau(n)+J-j)_+ + (n-j)\left(\sum_{k=0}^{\tau(n)+J} e^{S_{\tau(n)}-S_k}\right)^{-1}\right].\end{aligned}$$

Hence we get

$$\begin{aligned}\left(\sum_{k=0}^{n} e^{S_{\tau(n)}-S_k}\right)^{-1} &\le \frac{e^{S_j-S_{\tau(n)}}}{n-j}\,\mathbf{E}_w N_n(j) \\ &\le \frac{(\tau(n)+J-j)_+}{n-j} + \left(\sum_{k=0}^{\tau(n)+J} e^{S_{\tau\ (n)}-S_k}\right)^{-1}. \quad (54)\end{aligned}$$

Evidently, for $y > 0$

$$\begin{aligned}\mathbf{P}\left(\sum_{k=0}^{n} e^{S_{\tau(n)}-S_k} < y\right) &= \sum_{p=0}^{n}\mathbf{P}\left(\sum_{k=0}^{n} e^{S_{\tau(n)}-S_k} < y;\, \tau(n)=p\right) \\ &= \sum_{p=0}^{n}\mathbf{P}\left(\sum_{l=1}^{p} e^{S_l^-} + \sum_{r=0}^{n-p} e^{-S_r^+} < y;\, \widetilde{M}_p^- < 0,\, L_{n-p}^+ \ge 0\right). \quad (55)\end{aligned}$$

Further, note that from (26) and (27), as $\min\{p, n-p\} \to \infty$,

$$\mathbf{P}\left(\sum_{l=1}^{p} e^{S_l^-} + \sum_{r=0}^{n-p} e^{-S_r^+} < y \,\middle|\, \widetilde{M}_p^- < 0,\, L_{n-p}^+ \ge 0\right) \Rightarrow \mathbf{P}_{-,+}\big(\eta_1^- + \eta_2^+ < y\big). \quad (56)$$

If the condition (19) is met, then the generalized arcsine law holds true (see e.g. Theorems 8.9.9, 8.9.5 in [7]):

$$\lim_{n\to\infty}\mathbf{P}\left(\frac{\tau(n)}{n} \le x\right) = \frac{\sin\pi\rho}{\pi}\int_0^x t^{\rho-1}(1-t)^{-\rho}dt, \qquad x \in [0,1]. \quad (57)$$

Thus, for any $\varepsilon_1 > 0$ there exists a $\delta_1 \in (0, 1/2)$ such that

$$\mathbf{P}\left(\tau(n) \notin (n\delta_1, n(1-\delta_1))\right) \le \varepsilon_1, \quad (58)$$

which, combined with (55) and (56), shows that, as $n \to \infty$,

$$\mathbf{P}\left(\sum_{k=0}^{n} e^{S_{\tau(n)}-S_k} < y\right) \Rightarrow \mathbf{P}_{-,+}\big(\eta_1^- + \eta_2^+ < y\big). \quad (59)$$

A similar argument combined with (53) shows that

$$\mathbf{P}\left(\sum_{k=0}^{\tau(n)+J} e^{S_{\tau(n)}-S_k} < y\right) \Rightarrow \mathbf{P}_{-,+}\big(\eta_1^- + \eta_2^+ < y\big) \quad (60)$$

as first $n \to \infty$, and than $J \to \infty$. On the other hand, again using (57), we conclude that, within the range $j \in [\tau(n), n-1]$,

$$\begin{aligned}\frac{(\tau(n)+J-j)_+}{n-j} &\le I\{\tau(n)+J>j\}\frac{J}{n-j}I\big\{\tau(n)\ge n-\sqrt{n}\big\} \\ &\quad + I\{\tau(n)+J>j\}\frac{J}{\sqrt{n}-J}I\big\{\tau(n)<n-\sqrt{n}\big\} \\ &\le JI\big\{\tau(n)\ge n-\sqrt{n}\big\}+\frac{J}{\sqrt{n}-J}\xrightarrow{p}0 \qquad (61)\end{aligned}$$

as first $n \to \infty$, and than $J \to \infty$.

Using (59) and (60), (61) on the left- and right-hand sides of (54), respectively, proves (52) for $j \in [\tau(n), n-1]$.

For $\tau(n) - j > 0$ one can use similar arguments. The only difference is that in this case

$$(\tau(n)+J-j)_+ = \tau(n)+J-j,$$

and for $j < \tau(n)$, varying with n in such a way that $(\tau(n)-j)_+ = o(n)$, the conclusion (61) still holds by (58). Theorem 6 is proved. □

Acknowledgments. Research supported by the ARC Centre of Excellence for Mathematics and Statistics of Complex Systems. The second author was also supported by the Russian Foundation for Basic Research (grant 05-01-00035) and by the program "Contemporary Problems of Theoretical Mathematics" of the Russian Academy of Sciences. He is grateful to the Department of Mathematics and Statistics of the University of Melbourne for its hospitality while he was visiting the department.

References

[1] Afanasyev, V.I., Geiger, J., Kersting, G. and Vatutin, V.A. (2005) Criticality for branching processes in random environment. *Ann. Probab.* **33,** 645–673.

[2] Athreya, K.B. and Ney, P.E. (1972) *Branching Processes.* Springer, Berlin.

[3] Babilot, M., Bougerol, P. and Élie, L. (1997) The random difference equation $X_n = A_n X_{n-1} + B_n$ in the critical case. *Ann. Probab.* **25,** 478–493.

[4] Barabási, A.L. and Albert, R. (1999) Emergence of scaling in random networks. *Science* **286** (5439), 509–512.

[5] Biggins, J.D. and Grey, D.R. (1997) A note on the growth of random trees. *Statistics and Probability Letters* **32,** 339–342.

[6] Billingsley, P. (1968) *Convergence of probability measures.* Wiley, New York.

[7] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987) *Regular variation.* Cambridge University Press, Cambridge.

[8] Bollobás, B., Riordan, O., Spencer, J. and Tusnády, G. (2001) The degree of a scale-free random graph process. *Random Structure and Algorithms* **18,** 279–290.

[9] Borovkov, A.A. (1976) *Stochastic Processes in Queueing Theory.* Springer, New York.

[10] Borovkov, K.A. and Motyer, A. (2005) On the asymptotic benavior of a simple growing point process model. *Statistics and Probability Letters* **72,** 265–275.

[11] Devroye, L. (1988) Applications of the theory of records in the study of random trees. *Acta Informatica* **26,** 123–130.

[12] Doney, R.A. (1995) Spitzer's condition and ladder variables in random walks. *Probab. Theory and Related Fields* **101,** 577–580.

[13] Élie, L. (1982) Compartement asymptotic du noyau potentiel sur les groupes de Lie. *Ann. Scient. École Norm. Sup.* **15,** 257–364.

[14] Geiger J. and Kersting G. (2000) The survival probability of a critical branching process in a random environment. *Theory Probab. Appl.* **45,** 517–525.

[15] Hirano, K. (1998) Determination of the limiting coefficient for exponential functionals of random walks with positive drift. *J. Math. Sci. Univ. Tokyo* **5,** 299–332.

[16] Mahmoud, H. (1992) *Evolution of Random Search Trees.* Wiley, New York.

[17] Pittel, B. (1994) Note on the height of recursive trees and m-ary search trees. *Random Structure and Algorithms* **5,** 337–347.

[18] Rüschendorf, L., and Rachev, S. (1995) Probability metrics and recursive algorithms. *Adv. Appl. Prob* **27,** 770–799.

[19] Spitzer, F. (1964) *Principles of Random Walks.* Van Nostrand, Princeton, NJ.

[20] Vatutin, V.A. and Dyakonova, E.E. (2004) Galton-Watson branching processes in random environment. II: Finite-dimensional distributions. *Theory Probab. Appl.* **49,** 275–308.